International Research Journal in Global Engineering and Sciences. (IRJGES)
ISSN: 2456-172X | Vol. 4, No. 2, June 2019 - August 2019 | Pages 93-98

RECOGNITION AND CLASSIFICATION OF BRINJAL USING MACHINE LEARNING ALGORITHM

R. Maheswari[1], Reka [2]

Senior Assistant Professor, P.G Student, M.E(Communication Systems)
Department of Electronics and Communication Engineering, Agni College of Technology, Chennai

Abstract- The most common tropical vegetables in agriculture is Brinjal. It is an important process for separating different types of Brinjals in Agriculture. It is being cultivated in large numbers differing in size, shape and color. This is a major vegetable crop in several countries like the U.S.A, India, Japan, Indonesia, Italy, China, France, Bulgaria, and several African countries. Brinjal classification has been estimated to be 32 million tons production around the world. The most common methods that farmers use to identify the type of brinjal is by looking at the plant or vegetable and the prior experiences of the production of plants. Recognition and classification of Brinjal approach uses feature extraction of GLCM texture features and classification using Multi class SVM algorithm. The ability to identify the eggplant based on quality in the food industry which is the most important technology in the realization of automatic brinjal sorting machines in order to reduce the work of humans and also time consuming. This Algorithm helps in improving classification process and get better accuracy. It is essential to obtain the advanced image processing technique to classify the varieties of Brinjal.

Keywords- Brinjal Classification, Feature extraction, Gray-level-co-occurrence matrices (GLCM), Multi class Support vector machine (MLSVM)

1.0 Introduction

Brinjal (Solanum melongena) belongs to the family Solanaceae, is the most common, popular and principal vegetable crop grown widely for its shape, size and color of fruits. High productivity, wide adaptation and ease in availability make the crop find its place as a poor man's crop. It is a fairly good source of calcium, phosphorus, iron and vitamins particularly – B group. Eggplant has cholesterol reducing properties due to the presence of a higher amount of poly-unsaturate+d fatty acids in pulp and seeds. It is a warm season crop and requires a relatively long growing season with plenty of sunshine and moderate day temperature. It can be grown in all types of soils, but well drained fertile, rich in organic matter silt or clay loam soils are ideal. It is a versatile yield adapted to various agro-climatic regions and can be developed consistently. Various cultivars are grown in India, consumer inclination being dependent upon fruit color, size and shape.

BRINJAL VARIETIES:

Eggplant can be divisive. The notorious nightshade is both beloved and bemoaned for its spongy texture and tiny, bitter seeds. Brinjal export from India involves several processes including documentation, health and safety measures and so on. There are several varieties of Brinjals around the world. Here are some major varieties of Brinjals used in India. The Pusa Ankur brinjals are oval round in shape and small sized. The fruits are glossy and very attractive and weigh each around 60-80g. This variety of brinjals is an early bearing one. It is ready for first picking around 45 days after transplanting. These fruits do not lose color or tenderness even on delayed picking. The Manjri Gota variety of Brinjal generally has a dwarf and spreading growth habit. The fruits of this variety are medium large in size,

round in shape and purple coloured with white strips. Upon maturity, the brinjals attain a golden yellow color. The average yield of this variety is 15-20 t/ha.

2.0 Literature Survey

The **Classifications of citrus fruit using image processing – GLCM** parameters[3] is based on classification of selected citrus fruits like orange, sweet lime and lemon based on Gray level co-occurrence matrix. The approach used a single view fruit image i.e. one side of the fruit has been captured. The ISH color model is utilized and decision guidelines are derived from the hue color. This procedure is an extensive alternative and a more achievable method for grading large amounts of Chokun oranges contrasted with manual grading. It utilizes the chrominance and intensity data from natural outdoor scenes as a method of guidance for a robotic manipulator in the harvest of fruit. This classification model could segregate oranges from the natural background of an orange grove utilizing only color data in digital color images. The present paper reviews with special focus to citrus fruits, the evolution of traditional fruit sorting and grading techniques and subsequent machine vision-based system with improved fruit handling and improved techniques for inspecting the fruits for various quality parameters like size, shape, color, defect, etc.

The paper **Gender identification of imperfect flowers using image classification**[9] presented two approaches that can be used to classify images of imperfect flowers according to their gender based on their RGB values. Two approaches were used to classify images of imperfect flowers according to their gender. This method would be able to identify the petals, stamens, pistils and other objects in the images. From the original image, the image of stamen and pistil is extracted. Multilayer perceptron classifier is used as a hidden layer with 3 nodes and an output layer with 4 nodes and the threshold values are obtained from it. The algorithm uses the values to classify an unknown image of bitter gourd vines in real time without using complex data mining methods such as deep learning algorithms. When the images have not only the flowers but other objects that are found around a bitter gourd vine, the accuracy of classification using image filters and classifiers of WEKA reduced slightly.

The paper on **Pure-CNN: A framework for fruit images classification**[2] proposed a Pure Convolutional Neural Network with a minimum number of parameters to recognize multiple fruits more accurately. The proposed fruit recognition approach is based on Pure Convolutional Neural Network (PCNN) framework with GAP layer. PCNN consists of partially connected layers, two of them followed with stride 2 and a GAP layer. To increase the efficiency and reduce time, a proficient GPU was used in the experiment. To make comparison, classical CNN with dropout and without dropout is being used for classification on the same dataset. The input image is fed into the convolutional layer. Convolutional layer extract features maps by linear convolutional filters from input images followed by nonlinear activation functions. The extracted feature maps pass through another layer, two of them followed with stride 2. In PCNN, the convolutional layers along with stride for down sampling is used. Linear Rectifying Units are used for hidden layer in deep learning. It has output 0 if the input is less than 0 and if the input is greater than 0 the output will be equal to the input. GAP is used to minimize the number of parameters and to protect model from overfitting. It is similar to max pooling layer but it performs more extreme type of dimension reduction. The SoftMax layer which is the last layer is used for predicting

hundreds and thousands of classes. The output of SoftMax is equal to the total number of fruit classes. Using GAP layer has not only proved to give better accuracy for fruit image classification but also prevents whole structure from overfitting. In most cases, the approach obtained the highest classification accuracy.

A convolutional neural network algorithm to establish the **fruit fly feature automatic extraction classification system**[13]. It provides effective classification and solves the classification caused by the manual design features in the traditional classification method. The proposed model can automatically extract the features of the fruit fly pests for effective classification, and solve the classification caused by the manual design features in the traditional classification method. The cumbersome and complicated process of classification of fruit flies is simplified, and the work efficiency of quarantine personnel is improved. So, the model has a good application prospect. At the same time, the next goal is to classify fruit flies of complex background, such as natural background, multi object background, and the similar color background, and so on.

A genetic algorithm on **Use of color leaf images to identify nitrogen and potassium deficient tomatoes**[6]. The diagnosing system of tomato plant disease in nutrient deficiency based on computer vision is proposed in this paper. It also provides theory and technology support for future online supervision of nutrients supply and long-distance diagnosis of nutrient deficiency. Tomato disease of nutrient deficiency often occurs, and the need for a new efficiency method for diagnosis is urgent. Color and texture features of leaves are extracted by some methods such as percentage intensity histogram, percent differential histogram, Fourier transform and wavelet packet. Additionally, Genetic Algorithm (GA) has been utilized to select features to get the best data for diagnosing the disease. The spatial domain leaf images were converted into frequency domain using Fourier transform. The periodic features related to the symptoms of disease were analysed in frequency domain. According to the characteristic of nutrients deficiency symptoms multiangle color and texture features were extracted. Then GA was used to optimize the features combination. In addition, the establishment of binary tree classification framework greatly could cut down the difficulty in pattern recognition. Moreover, the diagnosis system could identify the nutrients deficiency 10 days before the exports could determine.

3.0 Proposed Method

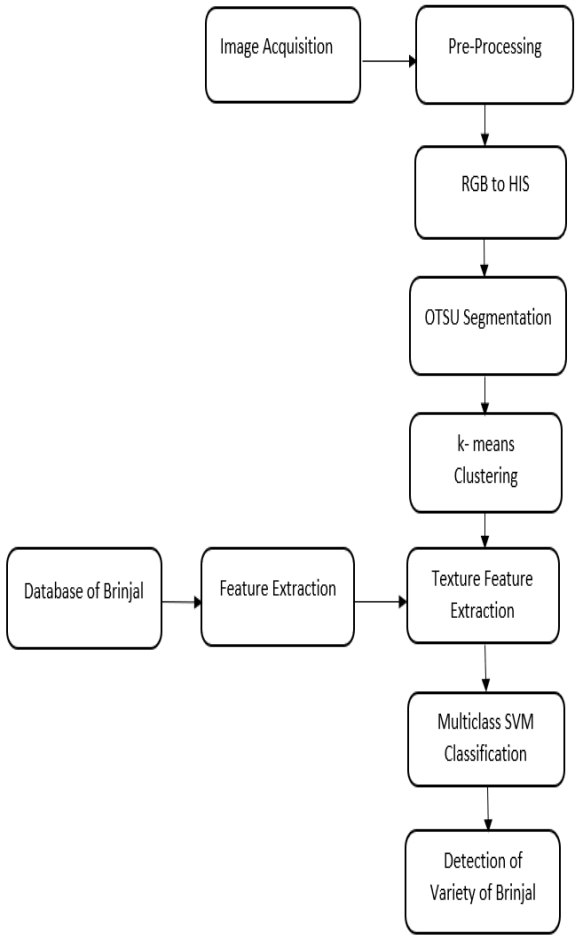

Fig 1 Flow chart of Proposed system

a. Preprocessing:

The manual cropping is done in order to get the region of interest in the brinjal images. The images after manual cropping are then resized. Some images, originally from cameras, manifest uneven lighting called shade. Due to variation in outdoor lighting conditions, some regions are brighter and some others are darker than the mean value for the whole image. This phenomenon is a consequence of inaccuracy in the system. Precise tuning of the camera is done to minimize this effect. The images contain some artifacts induced like scratched, coat or mark, lumps of dust or abrasive particles. Hence, median filters have been used to remove such artifacts.

Gray = 0.299 * R + 0.578 * G + 0.114 * B

Where R (Red), G (Green) and B (Blue) correspond to the color of the pixel.

b. RGB to HIS

The HIS model describes colors similarly to how the human eye tends to perceive color. RGB characterizes color in terms of a combination of basic colors. 'Hue' represents the color, 'Saturation' represents the amount to which that respective color is mixed with white and 'Intensity' represents the amount to which that respective color is mixed with black. Θ is the angle between each axis.

c. OTSU Segmentation

In general, autonomous segmentation is one of the most difficult tasks in digital image processing. Otsu's thresholding chooses the threshold to minimize the intraclass variance of the thresholder black and white pixels. It is a global thresholding technique that uses the histogram of the image for the threshold searching process. The segmentation will maximize "between class variance" of the segmented classes. Otsu proves that minimizing "within class variance" is the same as maximizing "between class variance" is computationally less expensive than "within class variance".

d. k-means clustering

Clustering is a technique to divide a set of information into a particular number of groups. One of the popular methods is k-means clustering. In k-means clustering, it partitions a collection of data into a k number group of data. It classifies a given set of data into k number of disjoint clusters as shown in *Fig 1*.

e. Feature Extraction

Feature extraction plays an important role for identification of objects. After segmentation the Texture features are extracted. Texture defines how the color, roughness and hardness is dispersed in the image, Before applying the GLCM method to input images, we must be sure that we can get texture features for eggplant only. So, we convert all pixels outside the leaf using mask to (NaN) "Not a Number", to be ignored during GLCM calculation. After that, GLCM has been calculated using the following angles (0◦, 45◦, 90◦, 135◦). So that, we have got 4 GLCM.

Multiclass classification should not be confused with multi-label classification, where multiple labels are to be anticipated for each instance. While some classification methods normally allow the utilization of multiple classes, others are by nature binary estimations; these can, notwithstanding, be transformed into multinomial classifiers by an assortment of strategies. This technique involves training a solitary classifier per class, with the samples of that class as positive samples and remaining samples as negatives. This procedure requires the base classifiers to make a genuine esteemed confidence score for its choice, instead of just a class name; discrete class labels alone can prompt ambiguities, where different classes are anticipated for a solitary sample. SVM is generally a two-class classifier. Practically, we have to handle problems involving k > 2 classes. Different strategies have been proposed for consolidating multiple two-class SVMs to construct a multiclass classifier. Generally used: one-versus rest approach.

4.0 Conclusion and Future scope

In this paper, the collected image is given as input and it is used for further processing. The input images may be in various size, different color combination and it contains noise are removed in the pre-processing section. After pre-processing, the k-means algorithm will be used to compute the distance from each centroid to points on a grid. It is also used to find groups which have not been explicitly labelled in the data. The appropriate cluster will be selected from the cluster of images and its texture features were extracted. Then the algorithm will be used to identify and classify brinjal varieties. Once the useful clusters are determined with the texture feature extraction the images will be processed with the SVM classification of brinjal with the trained Data set given to the algorithm.

References:

[1] Ashok Kanade, Arvind Shaligram, "Development of Machine Vision based system for classification of Guava fruits on the basis of CIE1931 chromaticity coordinates" 2nd International Symposium on Physics and Technology of Sensors, 8-10th March, 2015, Pune, India.

[2] Asia Kausar, Mohsin Sharif, JinHyuck Park and Dong Ryeol Shin, presented the paper on "Pure-CNN: A framework for fruit images classification", 2018 International Conference on Computational Science and Computational Intelligence (CSCI).

[3] Chandan Kumar, Siddharth Chauhan, R. Narmadha Alla and Harika Mounica gurram authors for "Classifications of citrus fruit using image processing – GLCM parameters", IEEE ICCSP 2015.

[4] Dr.S.M. Mukane1, Ms.J.A.Kendule, "Flower Classification Using Neural Network Based Image Processing", IOSR Journal of Electronics and Communication Engineering (IOSR-JECE).